



ECOSYSTEMS

The state of open source in Europe

Inside the key themes from this year's main open-source developers' event

April 11, 2024 - 12:00 pm



Open source is at a crossroads. For the past few years, venture capital has directly or indirectly paid for many of the contributors and much of the infrastructure it needed to keep going.

That was until the past 24 months or so, when funding started to slow down, leading to less internal development or funding resources going toward open source.

Skip to content



Latest Deep tech Sustainability Ecosystems Data and security Fintech and ecommerce Future of work Podcast

Earlier this year, I attended [Europe's](#) preeminent open-source event, [FOSDEM](#) (Free and Open source Software Developers European Meeting). Running for 24 years, the 2024 event attracted around 10,000 visitors and held 854 talks spread across Brussels' ULB. All for free and largely organised by a handful of organisers and sponsors.

TNW Conference 2024 - Calling all Scaleups to join on June 20-21

Showcase your scaleup in front of investors, change-makers and potential customers with our curated Scaleup package.

Register Now

In some ways, FOSDEM is a perfect microcosm of what open source should be, and what better place to analyse the current state of open source in Europe and the community at large.

FOSDEM is such a sprawling event. It covers not only the dozens of [dev](#) rooms on the university campus but also several side events around Europe (I also attended [State of Open Con](#), organised by Open UK in London). As such, it's hard to extract key themes, but in this article, I attempt to give an overview of those I saw and heard.

These are artificial intelligence (AI), the environmental impact of tech, project sustainability, governance and regulations, and the data that underpins all of these.

Open AI

There's a certain irony that whilst the arguably most famous [AI](#) company has "open" in its name, [it's decreasingly open in its research and development](#). Behind the scenes of the AI hype, many people in open source are concerned about how a few companies control most of the data that power many AI tools.

Many of us are used to the more traditional data sources behind applications, such as databases or API responses, where we can inspect their data easily. Models AI tools use, typically called Large Language Models or LLMs, represent their underlying data in a way that's much harder for a developer to see without additional tooling.

While much of the tooling around AI is open source, it doesn't mean the resulting models are, or even if they are, that it's possible to "see what's inside them." The open source community is working hard to change this, especially those who attend FOSDEM, where "open" is more of a way of life than an ideal.

As Niharika Singhal of the FSF for Europe mentioned in [her talk in the AI and Open Source dev room](#): "The popularisation of the (miss)use of the term 'open' in AI systems is particularly concerning."

She continued to mention that the open-source community needs to take projects to task that label themselves as "open" if they don't adhere to certain principles.

Open Source Initiative

Stefano Maffulli of [Open Source Initiative](#) (OSI) [took this topic further](#) to cover the initiative's collaborative definition that a truly open AI system needs to be available under legal terms that grant the freedoms to:

- **Use** the system for any purpose and without asking for permission.
- **Study** how the system works and inspect its components.
- **Modify** the system to change its recommendations, predictions or decisions to adapt to your needs.
- **Share** the system with or without modifications for any purpose.

Post-FOSDEM, the OSI continues to work on the definition. Stay up to date and get involved at <https://opensource.org/deepdive/drafts>.

OpenLLM Europe

Always a [bastion](#) of data sovereignty, FOSDEM saw the launch of [OpenLLM Europe](#), a Europe-wide community initiative to create the first **Skip to content** gual language model. This is to ensure that model content is more European-focused and includes a wider gamut of



[Latest](#) [Deep tech](#) [Sustainability](#) [Ecosystems](#) [Data and security](#) [Fintech and ecommerce](#) [Future of work](#) [Podcast](#)

tatives in several of the major European languages as well as smaller ones such as Maltese, Irish, and Slovak.

AI tooling

While it's always hard to pin down many open source projects to a particular region, on the more technical tooling side, most of the AI talks focussed on building, running, and training models "locally." Open doesn't mean easy, and working with models still requires significant time and resources.

Tying together the technical with the conceptual presented above, [Julie Hunter showed](#) how she built new models based on the French language and also re-tuned existing models, such as those from [Mistral AI](#) (ironically based in France), to answer more often in French than it currently does.

Hunter demonstrated one example asking for a recipe in French. It would start the response in French and halfway switch to English. The recipe was even originally a French dish.

A long-term player in European open source, [Nextcloud showcased their new local AI tools](#) that provide a user interface for finding and using models on a Nextcloud instance.

The features provide convenience wrappers and use a traffic light system to rate the models in the catalogue, bearing in mind many of the "open" aspects mentioned by Stefano Maffulli. I am unsure if my Raspberry Pi at home could cope with running any LLMs, but I look forward to trying.

Finally, an issue for many AI tools is getting meaningful and current data in and out of models in a spontaneous and dynamic way, for example, fetching daily news or conversations in a messaging app. Tuana Çelik from [Haystack](#) showcased an [open-source tool](#) to build components that use Retrieval-Augmented Generation (RAG) in a more "traditional" developer way to help you build AI data processing pipelines.

Sustainability

When I started my career in tech, compute resources were expensive. Servers, memory, and storage were so expensive that you had to find as many ways as possible to optimise how you used them. Then, with the advent of large multi-service cloud hosts, the costs dropped, and instead, we resolved many technical scaling issues by using more "cheap" resources.

Except they weren't actually as cheap as we thought. First, many of the new wave of cloud companies weren't necessarily cheaper; they just had different pricing models. The recent budgetary cutbacks in tech are causing many to question decisions made when money was more freely available.

However, this abundance of infrastructure also involves another cost. The environmental impact. "The internet," according to [the shift project](#), is estimated to have an impact roughly equal to that of the airline industry. Personally, I think this number is underrated, as the definition is vague and doesn't include a lot of the hardware and software that is connected to it.

Whether motivated by cost-cutting, forthcoming regulation, or pure altruism, monitoring and reducing the environmental impact of applications is now a big topic in open source, especially in Europe.

Monitoring environmental impact

The continent is home to several foundations and initiatives, including the [Green Web Foundation](#), whose [co2.js](#) is used by many to monitor their application's impact. [Mike Gifford gave an interesting talk](#) on efforts using [co2.js](#) to reduce the carbon impact of the most popular content management systems (CMS).

Among the waves of cool new technologies at FOSDEM, it's worth remembering that a handful of CMS, run the vast majority of the internet (WordPress alone runs about 60% of websites). As Mike said:

"It's easy to forget that real hardware is behind all the 'bits' we push around and use."

Most of the major browsers now have some form of carbon impact profiling built in, but head and shoulders above the rest is the [Firefox profiler](#) that goes into dizzying detail on the computing resources a web page is using.

[One of my favourite talks](#) was from Florian Quèze, one of the key engineers behind the tool. Estimating carbon emissions of applications is still in its nascent stages, and Florian detailed some of the engineering decisions taken to build calculations. For exam-

Skip to content



[Latest](#) [Deep tech](#) [Sustainability](#) [Ecosystems](#) [Data and security](#) [Fintech and ecommerce](#) [Future of work](#) [Podcast](#)

ways a surprise, it's fascinating to measure and see in such detail. Seeing the potential impact of an image or JavaScript framework visualised drives home what adding something that seems so small can mean.

Governance and funding

FOSDEM has plenty of content on the governance, funding, and broader sustainability of open-source projects. But, the topic was more the domain of [State of Open Con](#) (SOOCON), an event organised by OpenUK, Europe's only body for fostering relationships between open-source projects, government, and industry.

Making those connections directly on the subject of open AI, SOOCON held AI consultation sessions with [the UK's science and innovation department](#) and Home Office. OpenUK's Amanda Brok told me about the long queues outside each session and the fruitful discussions developers were bringing straight to the government.

"I think it's important that the authentic voice of the people who are actually doing the work is heard during the process. I don't think it has been elsewhere in the world, and I don't think it has been historically in the UK," Brok said.

Cost of cyber resilience

Europe is renowned for regulations, and the past year has resulted in several large policy frameworks that influence tech — [the Cyber Resilience Act](#) (CRA), [the Product Liability Directive](#), and the [EU AI Act](#). With lots of information to digest and react to, both FOSDEM and SOOCON held deep-dive sessions.

Over the past year, the CRA has been of the most concern to open-source communities, as it puts responsibility for harm caused by software into the hands of creators. For open-source software, this is complicated, as who is really responsible? The creator of the open-source software or its implementor? Many open-source projects have no legal entity that anyone can hold "responsible" for problems or harm.

After heated conversations between open-source communities and the EU last year, [EU lawmakers and regulators ended up listening and reacting](#) to most of the communities' feedback, but there are still concerns that it adds too much overhead to a new role the act calls "open-source stewards." At a time when funding is tight, [Tobie Langel asked the question](#) on everyone's mind:

"Who's going to bear the cost of this overhead?"

Granted, there are plenty of open-source foundations, but they also have limited resources and bring their own overhead to small projects.

The growing importance of data

Open source is [used more and in more places than ever before](#). However, due to financial pressure in the wider ecosystem and the influx of a new form of "open" projects, AI tools, and services, the community needs to be more considered and cautious than ever. Amanda Brok put it well when we spoke, telling me that SOOCON and OpenUK intentionally changed their stated aims a few months ago.

"'Open-source software and hardware' used to be enough to encompass the community and its aims and concerns. Now it's 'Open-source software, hardware, and data.'"

[An Open data movement](#), that aims to keep public data as freely accessible as possible, has existed for some time. [The EU alone has nearly 2 million data sets](#). However, this past year saw the open-source community have to care about the openness of data in completely new ways.

If a project or community wants to leverage new AI tools, it must now consider the data it uses and how it was trained. It needs to maintain regulatory data if it has any usage in the EU.

Depending on new regulations around the world, in the future, it might need to start profiling and reporting carbon impact. The wider commercial tech industry has known the value of data for some time, and now, whether it likes it or not, open source is also realising its importance and value.

Open to change

Despite a challenging 12-18 months in open source, the European ecosystem is in a good position to rise to the occasion. Compared to the US, open source in Europe has always been more conservative about funding sources and seeking different ways to sustain itself.

Skip to content



Latest Deep tech Sustainability Ecosystems Data and security Fintech and ecommerce Future of work Podcast

make application software and hardware more sustainable call Europe home. Thus, the continent is in an excellent position to lead the change.

There are challenges ahead, and open source needs to adapt, but I am confident that the European community is as ready as ever to face them.

STORY BY

Chris Ward

Technology writer, podcaster, and video maker by day. Fiction, games, and music by night. chrischinchilla.com

Get the TNW newsletter

Get the most important tech news in your inbox each week.

you@email.com

SUBSCRIBE

I agree to TNW storing and processing my personal data to receive the requested newsletter(s). For more information check out TNW's [Privacy Policy](#).*

Also tagged with

SOFTWARE DEVELOPER EUROPE OPEN SOURCE AI

Published April 11, 2024 - 12:00 pm UTC

[Back to top](#)

Skip to content



- Latest
- Deep tech
- Sustainability
- Ecosystems
- Data and security
- Fintech and ecommerce
- Future of work
- Podcast

[Search More Roles](#)

Related Articles



- Latest
 - Deep tech
 - Sustainability
 - Ecosystems
 - Data and security
 - Fintech and ecommerce
 - Future of work
 - Podcast
-

DEEP TECH

UCL spinout bags £10M to make AI 'super brains' for 100x faster LLM training

DEEP TECH

Italy sets up €1E

The heart of tech

More TNW

- Media
- Events
- Programs
- Spaces
- Newsletters
- Jobs in tech

About TNW

- Partner with us
- Jobs
- Terms & Conditions
- Cookie Statement
- Privacy Statement
- Editorial Policy
- Masthead

Skip to content



[Latest](#) [Deep tech](#) [Sustainability](#) [Ecosystems](#) [Data and security](#) [Fintech and ecommerce](#) [Future of work](#) [Podcast](#)
